

CLAIMS

We Claim:

5 1. A system for generating a corrected three-dimensional model of nucleic acids, comprising a processor configured to:

- a) generate an initial, uncorrected model of a test sequence by comparison to a reference sequence;
- b) align secondary structure constraints of a reference sequence with a test sequence to generate an aligned sequence;
- c) make substitutions, deletions, and insertions dictated by said aligned sequence using geometrical computation algorithms for said substitutions and using molecular mechanics and molecular dynamics algorithms to close gaps caused by said deletions and insertions;
- d) identify conserved hydrogen bonds present in both said reference sequence and said uncorrected model to select hydrogen bond constraints; and
- e) optimize said uncorrected model using a forcefield algorithm that accounts for said hydrogen bond constraints to generate a corrected three-dimensional model of said test sequence.

20 2. A method for generating a corrected three-dimensional model of nucleic acids, comprising the step of submitting a test sequence to the system of claim 1 under conditions such that a corrected three-dimensional model of said test sequence is generated.

25 3. A system for predicting nucleic acid three-dimensional structure, comprising a processor configured to:

- 30 a) compute a plurality of secondary structures of a test nucleic acid;

- b) decompose said secondary structures into nucleic acid structure motifs;
- c) rank said structure motifs in a hierachal tree;
- d) identify candidate three-dimensional motif structures for said motifs from a database of known three-dimensional structure motifs;
- 5 e) link said candidate three-dimensional motif structures in an order specified by said hierachal tree to generate a candidate three-dimensional composite structure;
- f) submit said candidate three-dimensional composite structure to an energy minimization algorithm to generate one or more refined candidate three-dimensional structure;
- 10 g. rank said one or more refined candidate three-dimensional structures based on the calculated total energy and optionally one or more scoring parameters comprising: solvent accessible surface area, molecular density, non-bonded energy; and
- 15 h) select a refined candidate three-dimensional structure based on best calculated energy to predict a three-dimensional structure of said test nucleic acid.

20

4. A method for generating a three-dimensional structure of a test nucleic acids, comprising the step of submitting a test sequence to the system of claim 3 under conditions such that a three-dimensional structure of said test sequence is generated.

25

5. A system for generating a nucleic acid structure motif database, comprising a processor configured to:

- a) receive nucleic acid physical structure information;
- b) decompose said physical structure information into nucleic acid structure motifs;

30

- c) associate data with said structure motifs, said data comprising: type of motif, size of motif, coordinates of backbone, and dihedral angles for bases;
- d) compare said nucleic acid structure motifs to existing motifs in said database; and
- e) add said structure motif and associate data to said database.

5

6. A method for generating a nucleic acid structure motif database, comprising the step of submitting nucleic acid physical structure information to the
10 system of claim 5.

7. A system for refining nucleic acid structure predictions, comprising a processor configured to:

- a) calculate energy minimization terms for a test nucleic acid structure prediction model, said energy minimization terms comprising: bond stretching, bond angles, torsion tress, and non-bonded interactions;
- b) optimize force constants, distance dependence, partial charges, and van der Waals radii parameters;
- c) account for gap penalties for insertions or deletions, if present in said prediction model;
- d) account for one or more experimental constraints associated with said test nucleic acid, said experimental constraints comprising hydrogen bonding information, position of phosphorus atoms, nuclear Overhauser effect information, residual dipolar coupling information, x-ray crystallographic electron density, cryo-electron microscopy information, and chemical probing information;
- e) employ distance constraints within a defined distance range but ignore distance constraints outside of said defined distance range; and

15

20

25

30

- g) account for one or more nucleic acid folding thermodynamic measures, said nucleic acid folding thermodynamic measures comprising: folding entropy, solvation entropy changes, enthalpy changes, and free energy changes.
- 5 h) account for known interactions, said known interactions comprising interactions with: proteins, metal ions, or other ligands by setting anchor points.

8. A method for refining a nucleic acid structure prediction, comprising the
10 step of submitting a nucleic acid structure prediction model to the system of claim 7.

9. A system for refining nucleic acid structure predictions, comprising a processor configured to:

- 15 a) generate an initial, uncorrected model of a test sequence by comparison to a reference sequence;
- b) align the reference sequence with the test sequence using secondary structure constraints to generate an aligned sequence;
- c) make substitutions, deletions, and insertions dictated by the aligned sequence using geometrical computation algorithms for the substitutions and using molecular algorithms to close gaps caused by the deletions and insertions;
- 20 d) automatically identify conserved hydrogen bonds present in both the reference sequence and the uncorrected model to select hydrogen bond constraints;
- e) allow input of other constraints determined by experimental or theoretical techniques, said other constraints comprising one or more of distances, dihedral angles, bond vectors, electron density, and accessibility to chemical modifying agents; and
- 25 f) optimize the uncorrected model using a forcefield algorithm to generate a corrected three-dimensional model of the test sequence that is consistent with selected and input constraints.

10. A method for generating a corrected three-dimensional model of nucleic acids, comprising the step of submitting a test sequence to the system of claim 9 under conditions such that a corrected three-dimensional model of said test sequence is
5 generated.